1

What This Book Is About

. . . qu'il faloit chercher dans la nature elle-même son Systême, s'il étoit vrai qu'ele en eût un . . . [. . .search in nature herself for her system, if it is true she really has one . . .]

(Michel Adanson 1763, p. clvij; translation in Nelson 1979, p. 21)

Is it not extraordinary that young taxonomists are trained like performing monkeys, almost wholly by imitation, and that in only the rarest cases are they given any instruction in taxonomic theory?

(Cain 1959, p. 243, quoted in Simpson 1961, p. vii and Felsenstein 1982, p. 379)

None whatever . . . the training I received in taxonomy was simply the attitude that this is purely incidental, everybody knows how to do it, there is no necessary theory, you need to know the rules of nomenclature, and that's all.

(Colin Patterson, in response to David Hull's question: 'Did you have any formal training in taxonomy?', unpublished taped interview)

This book is about biological taxonomy and classification as seen through the lens of cladistics. We set out to address the question: How is taxonomy done?

One way to find out how taxonomy is done is to examine any leading taxonomy journal. Inspection of any issue of the journal *Phytotaxa*, for example, is revealing. Issue 231(3) was published on 23 October 2015. It has 7 Articles and 3 Correspondence pieces, roughly 100 printed pages. Ignoring the Correspondence, the Articles deal with the taxonomy of a wide range of botanical groups (as understood in the older 'non-animal-photosynthetic' sense): flowering plants, fungi, diatoms, ferns and so on. The contributions are concerned with describing (diagnosing) new taxa (most often at the species level), revising groups of species (some are substantial taxonomic revisions), creating identification keys, making appropriate nomenclatural acts or various combinations of these endeavours.

It is a similar experience if a recent issue of *Zootaxa* is examined. Issue 4034 (1) was published on 28 October 2015. It has 9 Articles and 2 Correspondence pieces, roughly 200 printed pages. Likewise, these contributions are primarily concerned with the same issues as the authors contributing to *Phytotaxa* but applied to animals.

The 16 Articles from both journals are all beautifully illustrated, most are clearly written and, even if they are considered by some to be minor contributions to science, add something significant to our growing knowledge of life on this planet. In short, among these 300 pages, there is a wealth of new information about life on Earth.

Considering the contents of these contributions in more detail, only one gives any account of the method used to identify 'key' characters or the principles behind the characterisation of any taxa, new or otherwise. By this, we do not mean the process with which the authors record their observations on specimens to uncover their characters (homologues) – these are usually explained in great detail. We mean how they establish the definitions (diagnoses) of the taxa they subsequently describe.[1] The exceptions were authors who used molecular data, either in addition to or as the sole source of their evidence; these studies gave very precise details of how their data were analysed – but it is (almost) impossible to use DNA data without doing so. In each case, however, a different kind of analysis was applied, and none offered any specific justification for their choice – they simply chose one or another computer program (or some combination) to process their data.[2]

If a method was chosen, then most, if not all, appear to have emerged as a result of, or a reaction to, either the numerical taxonomy of the 1950s and 1960s, the various developments related to Hennig's *Phylogenetic Systematics* in the 1960s, 1970s and 1980s, the necessity of dealing with vast amounts of molecular data from the 1990s onwards – or, as is more often the case, following a tradition of what we will call 'empiricism', learnt from, or handed down, from previous generations' study of a particular group of organisms.

The central aim of this book, then, is to explore the method, or methods, in taxonomy, specifically methods of discovery – which we see as the aim of 'empiricism' – and how those results are presented in a classification.

How Is Taxonomy Done?

Bolstered by the age of bioinformatics, the dawning of cyber-taxonomy and the creation of mega-journals to cope with the vast amounts of taxonomic work that require rapid publication, unknown organisms are now being discovered, described and classified at an astonishing rate. One might, for example, consider the success of the mega-journal *Zootaxa* (Zhang 2006). Launched in 2001, by 2005

---

[1] Only one contribution assembled a data matrix and used PAUP to find a tree using parsimony for the basis of their classification (Tahseen & Mustaqim 2015)

[2] There were four contributions that used molecular data.

'*Zootaxa* [had] published 1103 issues in a total of 31 038 printed pages, with contributions from 1619 authors from around the world describing 2337 new taxa' (Zhang 2006). That is a period of four years – and that commentary is over a decade old. A more recent contribution – 'Describing unexplored biodiversity: *Zootaxa* in the International Year of Biodiversity' – noted:

> In the International Year of Biodiversity (2010), *Zootaxa* published 1,582 papers (including 92 monographs) in 405 issues, with a total of 32,330 pages. These papers included descriptions of 3,951 new taxa, of which 3,664 are of the species-group, 268 of the genus-group and 19 of the family-group. It is estimated that the total new animal species described in 2010 is most likely to be between 15,000 and 20,000, and *Zootaxa* has thus contributed 18 to 24% of the total. (Zhang 2011)

A massive achievement indeed: but one might scour those many pages and find little discussion concerning how taxonomy is actually done.

In some respects, however, the question is actually easy to answer. Consider this example: Imagine a dataset acquired from 100 representatives (specimens) each of a harvestman (e.g., *Wintonia scabra*, https://en.wikipedia.org/wiki/Opiliones), a huntsman (e.g., *Delena cancerides*, https://en.wikipedia.org/wiki/Delena_cancerides), a redback (e.g., *Latrodectus hasseltii*, https://en.wikipedia.org/wiki/Redback_spider) and a funnelweb spider (e.g., *Hadronyche versuta*, https://en.wikipedia.org/wiki/Hadronyche_versuta). From this dataset two characters are studied for each set of 100 specimens: (1) the presence of modified pedipalps (palps) and (2) spinnerets. The tarsus on the palp is developed in adult male spiders to transfer sperm. These characters are often referred to as homologues as they are names for parts of the organism, or more properly, the parts of the specimens (Chapter 7). Examining the funnelweb spider, the modified palp is different from those of the huntsman and the redback spider, and missing altogether in the harvestman. These data are informative as to their relationships: any particular funnelweb specimen is more closely related to all other funnelweb specimens than it is to any of the other organisms in this set. This detail can be written so:

W. scabra, *L. hasseltii*, *D. cancerides* (*H. versuta specimen #1*, *H. versuta specimen #2* . . . #100)

The round brackets group those that are more closely related to each other than to anything else. So in this example, all 100 specimens of *Hadronyche versuta* are represented by enclosing them within a pair of round brackets.

All modified palps found in spiders are considered to be the same organ, so a further statement can be made: all spiders are more closely related to each other, as they have modified palps, than they are to harvestmen, who do not. All palps are not necessarily the same, and modified versions of the character (homologue) are numbered individually in square brackets:

Character 1 (Ch. 1): no palps (palps [1], palps [2], palps [3])

This could be simplified as: Ch. 1: (0 (1, 2, 3)). Again, the rounded brackets group those items more closely related to each other than to anything else, in this case the set of homologues known as palps.

Taxon names can be substituted for the character:

W. scabra (*L. hasseltii*, *D. cancerides*, *H. versuta*)

The various types of spinnerets are also found in all spiders but not in harvestmen, yielding two characters with an identical set of relationships:

Character 2 (Ch. 2): no spinnerets (spinnerets [4], spinnerets [5], spinnerets [6])

Simplified as: Ch. 2: (0 (4, 5, 6))

Taxon names can be substituted for the character:

W. scabra (*L. hasseltii*, *D. cancerides*, *H. versuta*)

Two further characters can now be considered: the presence of (3) orthognaths (downward pointing fangs) and (4) labidognaths (inward pointing fangs). These yield the following relationship:

Character 3 (Ch. 3): no orthognath (orthognath [7], orthognath [8])

Simplified is: Ch. 3: (0 (7, 8))

Taxon names can be substituted for the character:

W. scabra, *L. hasseltii*, *D. cancerides* (*H. versuta*)

and

Character 4 (Ch. 4): no labidognath (labidognath [9], labidognath [10])

Simplified is: Ch. 4: (0 (9, 10))

Taxon names can be substituted for the character:

W. scabra, *H. versuta* (*L. hasseltii*, *D. cancerides*)

All four characters together yield the following:

W. scabra (*H. versuta* (*L. hasseltii*, *D. cancerides*))

This final statement of relationships among these animals is found merely by the addition of these 4 characters with 10 homologues ([1] – [10]):

Ch. 1: (0 (1, 2, 3)) + Ch. 2: (0 (4, 5, 6)) + Ch. 3: (0 (7, 8)) + Ch. 4: (0 (9, 10)) = [0] ([3, 6, 7, 8] ([1, 4, 9], [2, 5, 10]))

The taxon names can be substituted for the character, which may be appended to each name:

W. scabra^{0} (*H. versuta*$^{3\ 6\ 7\ 8}$ (*L. hasseltii1*$^{4\ 9}$, *D. cancerides*$^{2\ 10}$))

More data can be added in the form of a new specimen. For example, the trapdoor spider *Misgolas rapax* (https://en.wikipedia.org/wiki/Sydney_brown_trapdoor_spider) has three more homologues that are part of the palp [11], spinneret [12] and orthognathus fangs [13]. Added to the existing set of relationships reveals it to be more closely related to the funnelweb than to any other taxon:

W. scabra^{0} (*H. versuta*$^{3\ 6\ 7\ 8}$, *M. rapax*$^{11\ 12\ 13}$(*L. hasseltii*$^{1\ 4\ 9}$, *D. cancerides*$^{2\ 5\ 10}$))

This set of relationships conveys various details of the organisms themselves, the characters that define them and how closely related one is to another.

This particular notation above might appear a little cumbersome, so a simpler way would be to relate the whole as a written classification as follows:

- Arachnida
  - Assamidae
    - *Wintonia scabra*
  - Araneae (spiders)
    - Mygalomorphae
      - Hexathelidae
        - *Hadronyche versuta*
      - Idiopeae
        - *Misgolas rapax*
    - Araneomorphae
      - Theridiidae
        - *Latrodectus hasseltii*
      - Sparassidae
        - *Delena cancerides*

It should be easily appreciated that such a classification is inclusive, meaning that if another species of *Latrodectus* is found, say *Latrodectus antheratus*, then we would expect it to share features with other organisms in *Latrodectus*, Theridiidae, Araneomorphae, Araneae and so on.

Of course, when *Latrodectus antheratus* was found, a further interesting question can be addressed: *Latrodectus hasseltii* has only been found in Australia[3]; *Latrodectus antheratus* has only been found in Paraguay and Argentina. Why? And what of other species in the genus *Latrodectus*? Some estimate the number to be more than 30 (www.wsc.nmbe.ch/genus/3502).

Of course, with respect to features that define the taxonomic hierarchy, there is no certainty that all will be present in all organisms included in any one particular category. As William Bateson remarked:

> Nevertheless, if I may throw out a word of counsel to beginners, it is: Treasure your exceptions! When there are none, the work gets so dull that no one cares to carry it further. Keep them always uncovered and in sight. Exceptions are like the rough brickwork of a growing building which tells that there is more to come and shows where the next construction is to be. (Bateson (1908) 1928, p. 324)

Investigating further character systems, probing further the structure of organisms, searching for more specimens – in museums or in the field (Chapter 7) – is the stuff of taxonomy, and characters *are* found that do not correspond to known taxa, and known taxa *are* found to be not so clearly defined after all, and known taxa *are* found to be composed of unrelated organisms that need separating – this, then, is the key: *the scheme above is one of relationships, rather than of 'things' and the endeavours of taxonomists are to determine what exactly is and what exactly is not related in some specific way*.

One might read this account of the relationships among harvestmen and spiders and search for some particular theory or theories that guide, or have guided, the exploration. But there is none, save that observations (homologues) yield conclusions (taxa): what we understand as a version of *empiricism* is the version that probably all working taxonomists have adopted in one form or another, whether precisely stated or otherwise.

In terms of methodology, many books that deal with taxonomy never describe how one would go about discovering taxa, what that process might be, how the day to day activities of a taxonomist might work out. An exception is Richard E. Blackwelder's book simply entitled *Taxonomy* (1967) – but this is now a half a century old. To mention a few more recent books that specifically address taxonomy, *Describing Species* (Winston 1999) is a very readable and useful account of what one might need to tackle when undertaking a taxonomic study, with chapters on nomenclature, literature, material required, etc. Only at the very end of the book, in its final chapter, is there any mention of method at all, and this is found in a chapter entitled 'Further Studies in Systematics', as if discovery were achieved only *after* one learnt something about species, even genera, and then it is as if the

---

[3] Some 'alien' specimens have been found in Southeast Asia, Japan and New Zealand.

only kinds of explorations relate to phylogeny. *The New Taxonomy* (Wheeler 2008) is a collection of essays related to, or about, taxonomy, but it too lacks any chapter on how it is actually done; *Descriptive Taxonomy* (Watson et al. 2015), another excellent collection of essays relating to taxonomy, also has no guide as to method. One might read all these books with profit but still be left with one vital question: how is taxonomy done?

If the question is re-phrased as ‘how is classification done?’, then one inevitably encounters a central issue of the cladistic period: ‘It is informative that Hennig (1966), the father of phylogenetics, devoted about two thirds of his book to issues related to classification, and that resistance to his ideas centred on questions of classification (e.g. Mayr 1974)’ (Wheeler 2004, p. 576). More specifically, it centred on the influence cast by the *Modern Synthesis* (e.g., Huxley 1942; Mayr 1942) and its offspring *The New Systematics* (Huxley 1940), beginning in the 1940s and still with us today in one form or another (Pigiucci & Müller 2010[4]; see also the forward by Pigiucci & Müller in Huxley 2010). We tackle this period from the perspective of discovering relationships in Part II (Chapters 3–6).

To facilitate our discussion, definitions used in this book are briefly outlined in the next section. This is simply to provide early guidance to our usage as each term is explained and discussed more fully in Chapter 7.

Terms of Classification

Taxonomy is the study of classification (from the Greek τάξις *taxis*, meaning ‘arrangement’, and νομία *nomia*, meaning ‘method’). Biological classification and taxonomy are discussed further in Chapter 2.

Cladistics: The goal of cladistics is to discover natural classifications hence it might be prudent to define it as: *an approach that aims to find natural taxa within a natural classification*. This is the subject matter of the entire book.

Cladogram is a hierarchical branching diagram, which includes all species, fossil and Recent. This is discussed in detail in Chapter 7.

Cladistic parameter: All cladograms have a branching aspect. This is its cladistic parameter and is the parameter used to deal with any Natural Classification. This is discussed in detail in Chapter 7.

Monophyly and aphyly: Monophyly refers to those groups of organisms captured by the cladistic parameter; aphyly refers to those groups not so captured.

---

[4] Very little is said of taxonomy, systematics, classification or cladistics in this edited volume – what is noted is not particularly complementary (e.g, Callebaut 2010, p. 456). Although Koonin provides a broad view of a ‘postmodern synthesis of evolutionary biology’, it is a rather conventional approach to the subject of classification (Koonin 2011).

Trees: Cladograms are not *trees*. In the following chapters in most cases the word 'tree' will refer to a *phylogeny*, *phylogram* or *phylogenetic (phyletic) tree*. A tree is a special kind of cladogram. This is discussed in detail in Chapter 7.

Homologue: A homologue is a part of an organism, usually abstracted from a part of a specimen, or specimens. All organisms might be viewed as being made of parts, such that each organism can be broken into its structural components. Homologues, however, are not just parts – they are parts that imply relationships. Homologues are the primary evidence on which all subsequent taxonomy is based. Character and homologue are interchangeable, with the taxonomist using character more frequently.

Homology is the actual relation derived from the homologues.

Relationship: The representation of a particular kind of comparison, such that minimally of three things, two will be more closely related to each other than either are to the third. This statement is economically represented by A(BC). One might see BC as the *homologues*, and A(BC) as *homology*; the latter understood to be equivalent to *taxon* and *monophyly*, as well as the cladistic parameter.

Heterobathmy literally means 'different steps on a stair' (from the Greek *bathmos,* a step, stair or grade). A cumbersome word, but when applied to taxonomy, it translates into finding the correct 'step' on the taxonomic hierarchy for each character observed. That is, asking the question: where does this character fit?

How This Book Is Structured

This book is structured in a reasonably conventional way, composed of a number of chapters. Each has been written so that hopefully it will 'stand-alone' as a statement on the particular subject addressed. We have tried to avoid repetition, but some is inevitable. Each chapter has illustrative examples, taken either from our own work or from the literature. Alongside the cited references, each chapter includes a list of *Further Reading*. In some cases, we have provided commentary on our choices. The list of items is included for those who wish to pursue any particular chapter's subject in greater depth – or even to explore some ideas alternative to our own. Naturally, our recommendations (and comments) reflect our own bias. For this we crave your indulgence but hope that we have been broad-minded enough to cover a range of material. We have dispensed with the traditional end-of-book glossary. While we accept that a glossary provides neat, concise definitions, it usually excludes the author's reasons for adopting that particular use.

It is important for us to stress that we specifically did not want to produce a cookbook. If anything, we have attempted to produce a book that explains the ingredients, offers routes to sourcing the best products and makes suggestions for possible recipes. It is up to you, dear reader, to make your own meal.

References

Adanson, M. 1763. *Familles des Plantes.* Vincent, Paris.

Bateson, BD. 1928. *William Bateson, F. R. S., Naturalist: His Essays & Addresses, Together with a Short Account of His Life*. Cambridge University Press, Cambridge.

Blackwelder, RE. 1967. *Taxonomy*. John Wiley & Sons, New York.

Cain, AJ. 1959. The post-Linnaean development of taxonomy. *Proceedings of the Linnean Society of London* 170: 234–244.

Callebaut, W. 2010. The dialectics of dis/unity in the evolutionary synthesis and its extensions. In: Pigiucci, M & Müller, GB (eds), *Evolution: The Extended Synthesis.* MIT Press, Cambridge, MA, pp. 443–481.

Felsenstein, J. 1982. Numerical methods for inferring evolutionary trees. *Quarterly Review of Biology* 57: 379–404.

Huxley, JS. 1940. *The New Systematics*. Oxford University Press, Oxford.

Huxley, JS. 1942. *Evolution: The Modern Synthesis*. George Allen & Unwin, London.

Huxley, JS. 2010. *Evolution: The Modern Synthesis: The Definitive Edition*. Forward by Massimo Pigiucci & Gerd B. Müller. MIT Press, Cambridge, MA.

Koonin, EV. 2011. *Logic of Chance: The Nature and Origin of Biological Evolution*. FT Press, New Jersey. [Updated edition].

Mayr, E. 1942. *Systematics and the Origin of Species*. Columbia University Press, New York.

Mayr, E. 1974. Cladistic analysis or cladistic classification? *Zeitschrift für zoologische Systematik und Evolutionsforschung* 12: 95–128.

Nelson, GJ. 1979. Cladistic analysis and synthesis: principles and definitions, with a historical note on Adanson's *Familles des Plantes. Systematic Zoology* 28: 1–21.

Pigiucci, M. & Müller, G. (eds) 2010. *Evolution: The Extended Synthesis*. MIT Press, Cambridge, MA.

Simpson, GG. 1961. *Principles of Animal Taxonomy*. Columbia University Press, New York.

Tahseen, Q. & Mustaqim, M. 2015. A taxonomic review of the genus *Goffartia* Hirschmann, 1952 (Rhabditida: Diplogastridae) with a note on the relationship of congeners. *Zootaxa* 4034: 70–86.

Watson, MF., Lyal, C. & Pendry, CA. (eds) 2015. *Descriptive Taxonomy: The Foundation of Biodiversity Research.* Cambridge University Press, Cambridge, UK.

Wheeler, QD. 2004. Taxonomic triage and the poverty of phylogeny. *Philosophical Transactions of the Royal Society of London, Series B,* 359: 571–583.

Wheeler, QD. (ed.) 2008. *The New Taxonomy*. Systematics Association Special Volume Series. CRC Press, Boca Raton, FL.

Winston, JE. 1999. *Describing Species, Practical Taxonomic Procedure for Biologists*. Columbia University Press, New York.

Zhang, Z-Q. 2006. The making of a mega-journal in taxonomy. *Zootaxa* 1358: 67–68.

Zhang, Z-Q. 2011. Describing unexplored biodiversity: *Zootaxa* in the International Year of Biodiversity. *Zootaxa* 2768: 1–4.

Further Reading

Blackwelder, RE. 1967. *Taxonomy*. John Wiley & Sons, New York.

A little outdated but useful for the discussion of what was called 'Omnispective taxonomy', an attempt to be wholly empirical. The book's merits are in the critique of the then prevailing views of Mayr. '*The Omnispective System* . . . [is the] modern form of classical taxonomy, [and] has been in use for two hundred years and has produced 99% of the revisionary and monographic work in taxonomy' (Blackwelder, RE. 1977. Twenty five years of taxonomy. *Systematic Zoology* 26: 107–137).

Michel, E. (ed.) 2016. Anchoring biodiversity information: From Sherborn to the 21st century and beyond. *Zookeys* 550.

While focusing on Charles Davies Sherborn (1861–1942), 'The papers in this volume fall into three general areas. In the first section, seven papers present different facets of Sherborn as a man, scientist and bibliographer, and describe the historical context for taxonomic indexing from the 19th century to today. In the second section, five papers (with a major appendix) discuss current tools and innovations for bringing legacy information into the modern age. The final section, with three papers, tackles the future of biological nomenclature, including innovative publishing models and the changing tools and sociology needed for communicating taxonomy'.

Polaszek, A. 2010. *Systema Naturae 250 – the Linnaean Ark*. CRC Press, Boca Raton, FL.

Watson, MF., Lyal, C. & Pendry, CA. (eds) 2015. *Descriptive Taxonomy: The Foundation of Biodiversity Research*. Cambridge University Press, Cambridge, UK.

Wheeler, QD. (ed.) 2008. *The New Taxonomy*. Systematics Association Special Volume Series. CRC Press, Boca Raton, FL.

Polaszek, Watson et al. and Wheeler are all useful collections of papers on various aspects of the taxonomic enterprise, although, as we noted above, not much space is devoted to how taxonomy is actually done.

Winston, JE. 1999. *Describing Species, Practical Taxonomic Procedure for Biologists*. Columbia University Press, New York.

A useful book focused on the practice of species-level taxonomy.